POLS/CSSS 512 · Time Series and Panel Data for the Social Sciences

NOTES ON DIFFERENCE-IN-DIFFERENCE MODELS

Christopher Adolph

Department of Political Science

and

Center for Statistics and the Social Sciences

University of Washington, Seattle

Difference-in-difference methods for panel data

Suppose you wish to study the causal effect of some intervention on an outcome y and have only observational data, not an experiment

Let's assume the intervention is a binary treatment, $D \in \{0, 1\}$

Treatment could be simultaneous for all treated units
(and never happens at all to untreated units)

Or treatment could be staggered in time for different units
(some or all of which could end up treated at some time)

You might find or construct a longitudinal dataset with N units and T time periods, such that there are periods observed before and after treatment for each unit

Goal: credibly estimate the treatment effect by comparing the trend in y over time between the treated and untreated cases (hence, difference-in-difference)

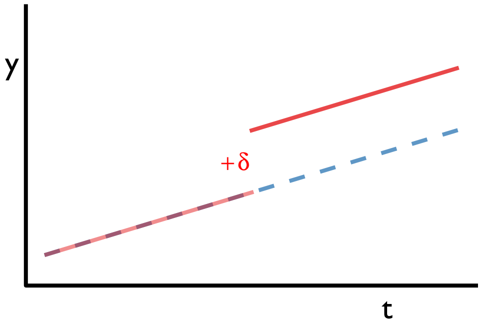
DiD model 1: assume immediate effects, parallel linear trends

$$y_{it} = \alpha_i + \gamma_1 t + \delta D_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

D_{it} is the treatment, coded as 1 in the period it takes place (and thereafter), and as 0 otherwise

The model includes unit fixed effects α_i and other time-varying controls $\mathbf{x}_{it}$

Untreated units follow the linear trend shown in blue; Once treated, units follow the same trend, but shifted by δ units



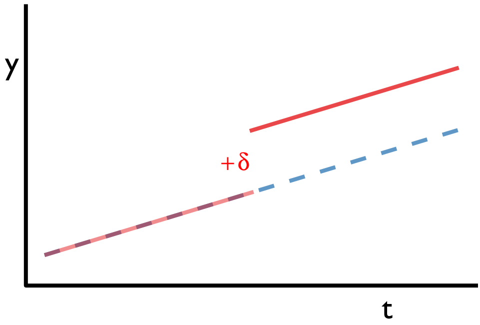
DiD model 1: assume immediate effects, parallel linear trends

$$y_{it} = \alpha_i + \gamma_1 t + \delta D_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

Easy to estimate with linear regression (e.g., `plm` in `R`)

Just a matter of correct specification and variable construction

Model 1: just add a treatment variable and a trend variable to a model with unit (individual) fixed effects

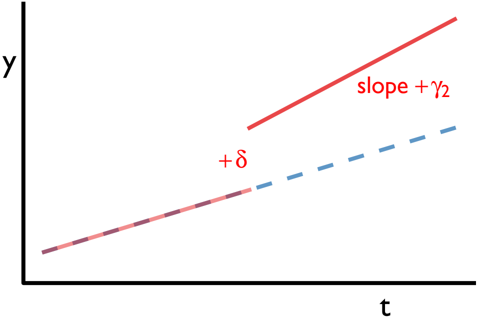


DiD model 2: change in intercept and linear trend

$$y_{it} = \alpha_i + \gamma_1 t + \delta D_{it} + \gamma_2 s_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

where $s_{it} \in \{0, T\}$ is the count of periods since $D_{i.}$ flipped from 0 to 1

Untreated units follow the linear trend shown in blue; Once treated, units follow a new trend, which is also shifted by δ units



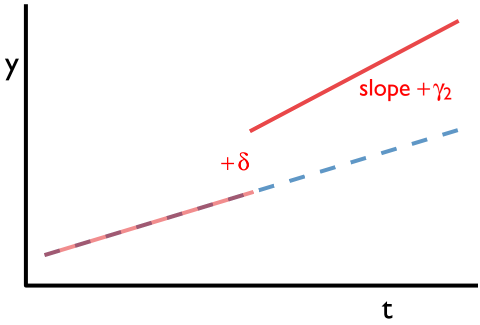
DiD model 2: change in intercept and linear trend

$$y_{it} = \alpha_i + \gamma_1 t + \delta D_{it} + \gamma_2 s_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

Easy to estimate with linear regression (e.g., `plm` in `R`)

Just a matter of correct specification and variable construction

Model 2: carefully create s_{it}, then add it, a treatment variable, and a trend variable to a model with unit (individual) fixed effects

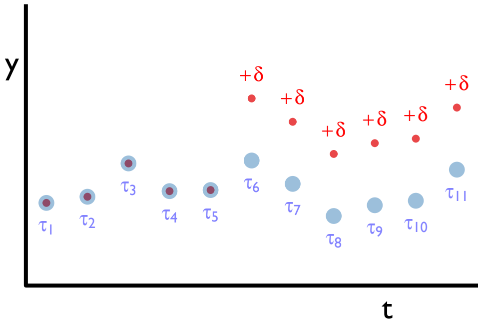


DiD model 3: only immediate change, but flexible parallel trends

$$y_{it} = \alpha_i + \tau_t + \delta D_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

τ_t is a period fixed effect, making this a "two-way" FE model; only feasible if some units are untreated and/or the treatment is (as-if randomly) staggered in time

Untreated units follow the flexible trend shown in blue; Once treated, units follow the same trend, but shifted by δ units



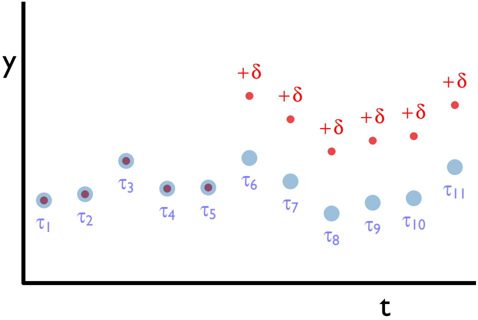
DiD model 3: only immediate change, but flexible parallel trends

$$y_{it} = \alpha_i + \tau_t + \delta D_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

Easy to estimate with linear regression (e.g., `plm` in `R`)

Just a matter of correct specification and variable construction

Model 3: add a treatment variable to a two-way fixed effects model



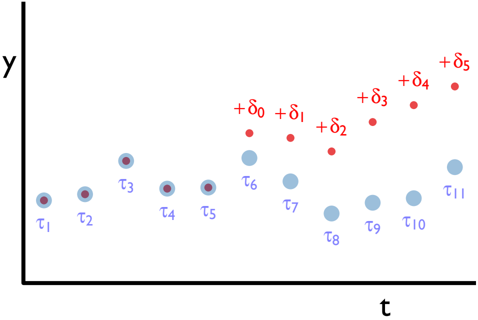
DiD model 4: Treatment changes flexible trends

$$y_{it} = \alpha_i + \tau_t + \delta_{t-t_i^*+1} D_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

where t_i^* indicates the period in which $D_{i.}$ flipped from 0 to 1, so $\delta_{t-t_i^*+1}$ indicates the treatment effect after $t - t_i^* + 1$ periods

Two-way fixed effects model with cumulative treatment effects

Untreated units follow the flexible trend shown in blue; Once treated, units follow a new trend that adds the (evolving) treatment effect to the period effect



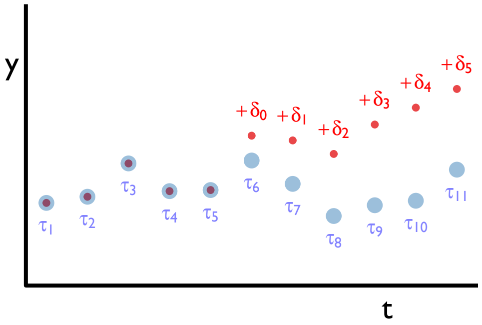
DiD model 4: Treatment changes flexible trends

$$y_{it} = \alpha_i + \tau_t + \delta_{t-t_i^*+1} D_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

Easy to estimate with linear regression (e.g., `plm` in `R`)

Just a matter of correct specification and variable construction

Model 4: Create a vector of dummy variables for each period from the treatment onwards, i.e, in $k \in t_i^*, T$; add these dummies to a TWFE model



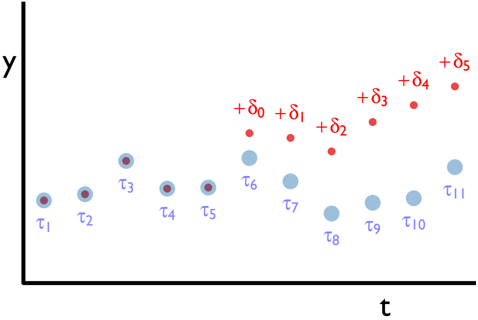
DiD model 4: Treatment changes flexible trends

$$y_{it} = \alpha_i + \tau_t + \delta_{t-t_i^*+1} D_{it} + \mathbf{x}_{it}\beta + \epsilon_{it}$$

Easy to estimate with linear regression (e.g., `plm` in `R`)

Be careful! If treatment is staggered, and very few (or just one) units are treated in the earliest treatment period, then the later $\delta_{t-t_i^*+1}$ may be poorly estimated or unidentified!

(In which case, consider splines in place of $\delta_{t-t_i^*+1}$, but there are other potential problems...)



DiD assumptions

In the absence of randomized experiments for many real world policy decisions, difference-in-difference models have become very popular in economics, political science, and policy evaluation generally

In ecomomics, variants of DiD now the modal methodology appearing in published quantitative papers

Under ideal conditions, simple two-way fixed effects models with a treatment variable *are* causally-identified DiD models.

However, causal identification rests on several assumptions, some of which are subtle...

DiD assumptions

However, causal identification rests on several assumptions, some of which are subtle:

Treatment is as-if random conditional on observables. Obviously the key assumption behind this model, so time- and unit-varying confounders still must be measured and controlled, and remain debatable in ways experiments are not.

Parallel trends. Units follow the same (either linear or flexible) trend up to the time of treatment.

Intuition check: how would you test these in Model 4?

DiD assumptions

However, causal identification rests on several assumptions, some of which are subtle:

Treatment is as-if random conditional on observables. Obviously the key assumption behind this model, so time- and unit-varying confounders still must be measured and controlled, and remain debatable in ways experiments are not.

Parallel trends. Units follow the same (either linear or flexible) trend up to the time of treatment.

Intuition check: how would you test these in Model 4?

$\rightarrow$ Can't be tested at the point of treatment, but can be tested in pre-treatment periods.

Test whether $\delta_{-1}, \delta_{-2}, \ldots$ are different from zero.

Treatment effect is homogenous.

If staggered, treatment time is random.

DiD assumptions & problems with them

Problem 1: Treatment may be endogenous. E.g., in the cigarettes example, states with (unobservable) stronger social interest in health may be more likely to choose treatment

Two-way fixed effects alone do not solve this problem if social interest in health is both time and unit varying (which it surely is)

Problem 2: Treatment effects may be heterogeneous, with units expecting larger effects opting for treatment sooner.

Surprisingly, problem 2 was only recognized recently. Can cause substantial bias in two-way fixed effects models

$\rightarrow$ Flurry of recent work on a variety of DiD methods to relax TWFE assumptions

A sample of recent innovations in DiD methods

Methods for heterogeneous treatment effects over staggered periods

Brantly Callaway and Pedro H.C. Sant'Anna (2021), "Difference-in-differences with multiple time periods," *Journal of Econometrics* 225(2), 200–230.

Clément de Chaisemartin and Xavier D'Haultfœuille (2020). "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review* 110 (9): 2964–96.

Kirill Borusyak, Xavier Jaravel, and Jann Spiess (2024)."Revisiting Event-Study Designs: Robust and Efficient Estimation," *Review of Economic Studies* 91 (6): 3253–3285.

Synthetic control methods.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010), "Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco Control Program," *Journal of the American Statistical Association* 105(490): 493–505.

Yiqing Xu (2017), "Generalized synthetic control method: Causal inference with interactive fixed effects models," *Political Analysis* 25(1): 57–76.

DiD for multiple post-treatment periods (Callaway & Sant'Anna, 2021)

Brantly Callaway and Pedro H.C. Sant'Anna (2021), "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 225(2), 200–230.

Brantly Callaway and Pedro H.C. Sant'Anna (2023), "Introduction to DiD with Multiple Time Periods," `https://bcallaway11.github.io/did/articles/multi-period-did.html`

Key insight: TWFE is biased when treatment effects are heterogeneous and $T > 2$.

Why? TWFE models implicitly compare treated units to a weighted average of (1) never-treated units, (2) units treated later, and (3) already treated units.

From a potential outcomes perspective, comparing treated units to (1) and (2) is good: you are comparing post-treatment to pre-treatment paths

Comparing to (3) compares two stages of post-treatment to each other!
This is can cause serious problems (even sign errors)

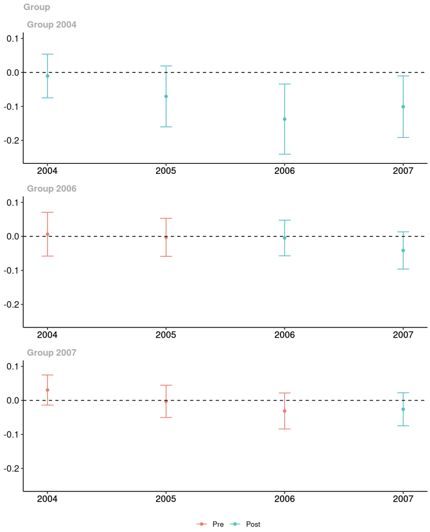
Solution: build an estimator around comparisons (1) and (2) only

DiD for multiple post-treatment periods (Callaway & Sant'Anna, 2021)

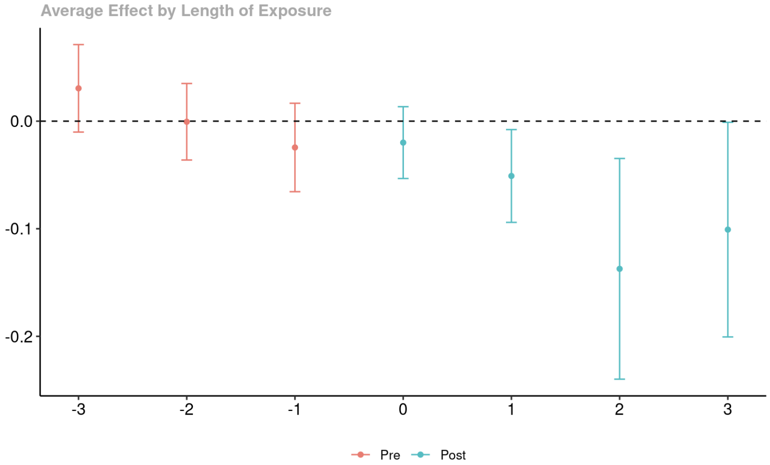
Brantly Callaway and Pedro H.C. Sant'Anna (2021), "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 225(2), 200–230.

Implemented in the R package `did` as `att_gt()`

Estimates *group-wise* average treatment effects for "groups" of units which differ in the time of treatment



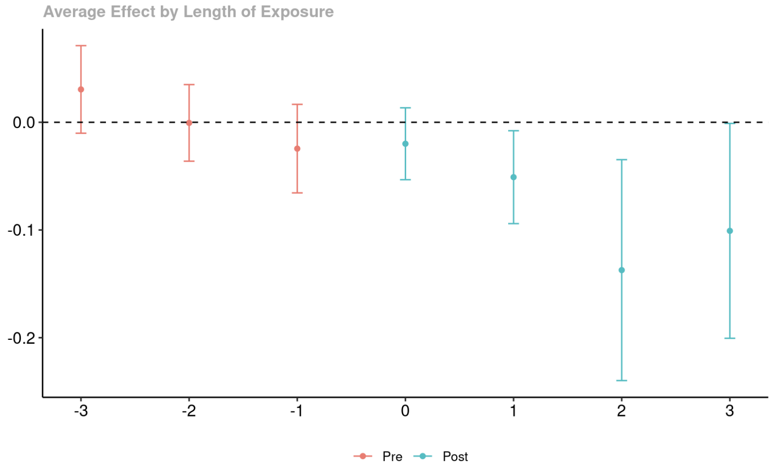
DiD for multiple post-treatment periods (Callaway & Sant'Anna, 2021)



Also gives an aggregate estimate of treatment effects *k* periods after treatment across groups

Usually the quantity of interest, but if heterogeneity is high (and especially if effect is decreasing over time), can you rely on the treatment effect in future cases?

DiD for multiple post-treatment periods (Callaway & Sant'Anna, 2021)



Note: the results in red to the left are tests of the parallel trends assumption

Intuition check: Why is that the case? What do they show?

Synthetic control methods

Basic idea is similar to matching

For each treated unit, construct a synthetic control unit from a weighted average of observed control units

Weights are chosen so that the synthetic control matches the treated case on pretreatment outcomes and covariates

But the weights then imply an outcome trajectory which the treated unit might have followed had it not been treated

The causal effect for the kth period post-treatment is just the difference between the observed outcome and the synthetic control in period k

$\rightarrow$ covariates enter through the construction of a "control unit," not as "control variables"

Can even be used when you have only one treated unit!

Synthetic control methods

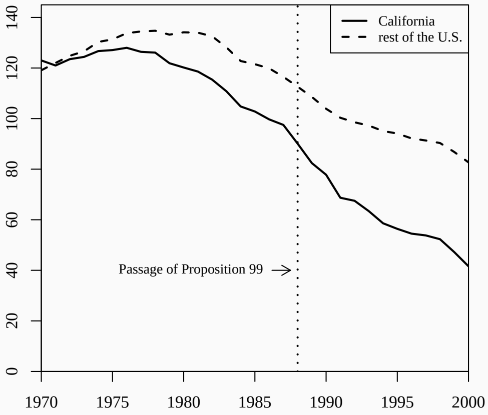
Consider an example from Abadie, Diamond, and Hainmueller (2010), who used a synthetic control to estimate the effect of California's Proposition 99, a 1988 law discouraging smoking

Rather than try to estimate a DiD model with only one treated case,
they build a synthetic California by weighting the average of 38 potential control states

	California		Average of
Variables	Real	Synthetic	38 control states
Ln(GDP per capita)	10.08	9.86	9.86
Percent aged 15-24	17.40	17.40	17.29
Retail price	89.42	89.41	87.27
Beer consumption per capita	24.28	24.20	23.75
Cigarette sales per capita 1988	90.10	91.62	114.20
Cigarette sales per capita 1980	120.20	120.43	136.58
Cigarette sales per capita 1975	127.10	126.99	132.81

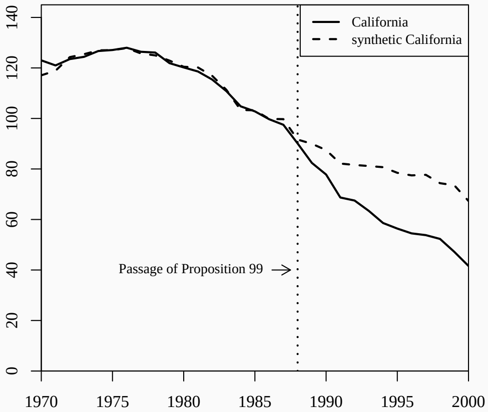
Note: All variables except lagged cigarette sales are averaged for the 1980-1988 period (beer consumption is averaged 1984-1988).

Synthetic control methods



Cigarette Consumption: CA and the Rest of the U.S.

Synthetic control methods



Cigarette Consumption: CA and Synthetic CA

Synthetic control methods

Above example can be replicated in the `synth` package

More general estimators exist for multiple treated units, e.g. Xu's `gsynth` package

More concern with having long pre-treatment time series to match on

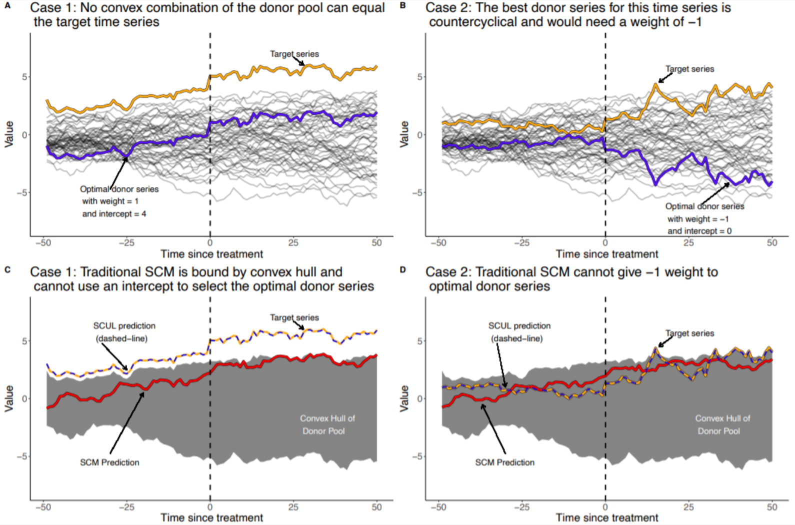
Less concern with the number of covariates (because they are being used to achieve balance through weights, rather than to estimate parameters)

Works best when you have some control units that are quite similar to the treated unit(s): similar to balance issues in any matching estimator

In particular, will not work if it needs to "extrapolate" from the convex hull of the control units to match a treated unit (something DiD doesn't require)

Synthetic control methods

Hollingsworth and Wing (2020)



DiD: Summing up

Overall, there is an explosion of alternative estimators for DiD in the last decade

There is easily enough material for a course (or more);
even keeping up with this literature is practically a full time job

Advice for a applied researcher:

(1) Consider which problems *and* opportunities your proposed panel analysis raises

(2) Consider which methods seem ideal, though potentially demanding of the data: try it

(3) Consider which other methods seem and to address concerns not dealt with by (2);
if appropriate and feasible, try it

(4) Consider which simple and familiar approaches are particularly tractable and relevant;
try them

(5) To the extent the above all agrees in findings; great
$\rightarrow$ publish the easiest to explain result, put the rest in a robustness section

If they disagree, report all you can, and use your expertise to adjudicate